

glmnetUI : glmnet (Lasso 與彈性網絡) 套件的互動式介面

繁體中文譯本，由人工智慧輔助翻譯並經編輯部審校；英文版為權威版本。

William Bert Craytor

第 1 卷，第 1 期 · 2026 年夏季

發表日期：2026 年 7 月 1 日

DOI: [10.67330/ja8s2k39](https://doi.org/10.67330/ja8s2k39)

© 2026 William Bert Craytor. Published by Pacific Vista Net.

本開放取用文章依據 [創用 CC 姓名標示 4.0 國際授權條款 \(CC BY 4.0\)](#) 的條款發布。

摘要。 glmnetUI 是 R 語言 glmnet 套件的圖形使用者介面 (graphical user interface)，該套件透過 Lasso、嶺迴歸與彈性網絡 (elastic net) 罰則來擬合正則化的廣義線性模型 (generalized linear model)。它提供三種用途模式——通用預測建模、不動產估價與市場區域分析——並引導使用者完成資料匯入、模型設定、交叉驗證擬合、係數路徑檢視以及報告下載。本文記錄了 glmnetUI 的資料格式要求、建模工作流程、輸出顯示介面以及完整的功能參考。

關鍵詞： Lasso，嶺迴歸，彈性網絡，正則化迴歸，glmnet，R，圖形使用者介面，不動產估價

出版者按 (2026 年 7 月)：付印前不久，glmnetUI 已併入 earthUI (0.11.0 及以後版本)，本文所述的彈性網絡工作流程現與 MARS 和 GAM 工作流程一同由 earthUI 提供。glmnetUI 仍以其最終發布形態 (0.5.0) 保留在 <https://github.com/wcraytor/glmnetUI>；後續開發將在 earthUI 中繼續。本文適用於已發布的 glmnetUI。

目錄

1	引言	3
1.1	歷史與背景	3
1.1.1	Lasso	3
1.1.2	嶺迴歸與彈性網絡	4
1.1.3	glmnet 套件	4
1.1.4	與 MARS (earth) 和 GAM (mgcv) 的比較	4
1.1.5	建議的工作流程	5
1.2	什麼是 glmnetUI ?	5
1.3	三種用途模式	5
1.4	不動產專用功能	6
1.5	快速上手	6
2	MLS 輸入資料要求	7
2.1	檔案格式與結構	7
2.2	估價模式所需的欄位	7
2.3	特殊欄位命名慣例	8
2.4	資料品質與完整性	8
2.5	標的列的位置	8
3	通用模式	9
3.1	概述	9
3.1.1	跳過第一列	9
3.1.2	按用途分別儲存設定	9

3.2	側邊欄工作流程	9
3.3	專案	10
3.4	從 earthUI 匯入 (選用)	10
3.5	主面板分頁	10
3.6	設定的保存	11
3.7	深色模式	11
3.8	地區與區域設定	11
3.8.1	儲存預設值	12
4	估價模式	13
4.1	標的列的處理	13
4.2	基準日與成交時距	13
4.3	特殊欄位指定	13
4.4	RCA 調整概覽	13
5	市場區域分析模式	14
5.1	與估價模式的差異	14
5.2	何時使用市場模式	14
6	變數選擇	15
6.1	目標變數	15
6.2	預測變數表	15
6.3	資料型別偵測與手動覆寫	15
6.4	預期符號與約束的施加	15
6.5	特殊欄位型別參考	16
7	參數選擇	17
7.1	Alpha (混合參數)	17
7.2	Lambda 的選擇	17
7.3	Family (分布族)	17
7.4	Standardize (標準化)	17
7.5	符號約束	18
7.6	鬆弛 Lasso	18
7.7	隨機種子	18
7.8	交互項矩陣	18
7.9	禁止進入主效應 (僅作交互項)	19
7.10	進階參數	19
7.11	設定預設值	19
8	擬合模型	20
8.1	擬合按鈕	20
8.2	擬合過程中發生了什麼	20
9	結果分頁	21
9.1	Data Preview (資料預覽)	21
9.2	Equation (方程式)	21
9.3	Correlation (相關性)	21
9.4	Summary (摘要)	21
9.5	Coefficients (係數)	21
9.6	Variable Importance (變數重要性)	22
9.7	Contributions (貢獻)	22
9.8	ANOVA (變異數分析)	22

9.9	Diagnostics (診斷)	22
9.10	Glmnet Output (Glmnet 輸出)	22
9.11	Report (報告)	23
10	下載資料	24
10.1	輸出欄位	24
10.2	CQA 分數	24
11	RCA 計算與下載	25
11.1	開啟 RCA 對話方塊	25
11.2	CQA 分數內插	25
11.3	輸出欄位	25
12	銷售比較表格	26
12.1	可比案例選擇對話方塊	26
12.2	活頁簿版面	26
12.3	工作表保護與剩餘特徵儲存格	26
13	下載報告	27
13.1	兩階段 Quarto 工作流程	27
13.2	報告格式	27
13.3	報告內容	27
14	與 earthUI 的比較	29
14.1	主要差異	29
14.2	共有的功能	29
14.3	何時該用哪一個	29
15	示範資料集：Appraisal_1.csv	30
15.1	說明	30
15.2	各欄位	30
15.3	建議的快速上手流程	30
16	系統需求與疑難排解	32
16.1	支援的平台	32
16.2	選用的相依套件	32
16.3	各平台說明	32
16.4	功能降級	32
16.5	疑難排解	32

1 引言

1.1 歷史與背景

1.1.1 Lasso

Lasso (Least Absolute Shrinkage and Selection Operator, 最小絕對收縮與選擇運算子) 由 Robert Tibshirani 於 1996 年提出。這個名稱是對牛仔套索 (lasso) 的有意雙關——套索是一種既能約束又能選擇的繩索。Tibshirani 之所以選用這個名字，是因為該方法真的會「套住」係數，約束它們的絕對值，並把其中一部分拉到恰好為零，從而選出哪些變數留在模型之中。

Tibshirani 的洞見在於給最小平方目標函數加上一個 L_1 罰則 (penalty, 係數絕對值之和)。嶺迴歸 (ridge regression, Hoerl 與 Kennard, 1970) 所用的 L_2 罰則只會把係數向零收縮而永

遠遠不到零；與之不同， L_1 罰則產生的是稀疏解，其中許多係數恰好等於零。這使得 Lasso 既是一種正則化（regularization）方法，也是一種變數選擇方法。

1.1.2 嶺迴歸與彈性網絡

嶺迴歸（ridge regression）由 Arthur Hoerl 與 Robert Kennard 於 1970 年提出。它添加的是 L_2 罰則（係數平方之和），會按比例收縮所有係數，但把所有變數都保留在模型中。當預測變數之間存在相關性（多重共線性）時，嶺迴歸很有效，但它不會簡化模型。

彈性網絡（elastic net）由 Hui Zou 與 Trevor Hastie 於 2005 年提出，它把兩種罰則結合起來。混合參數 α 控制兩者的配比： $\alpha = 1$ 為純 Lasso， $\alpha = 0$ 為純嶺迴歸，介於其間的取值則給出折衷方案。彈性網絡克服了 Lasso 的一個侷限：當預測變數高度相關時，Lasso 往往只選中其中一個而忽略其餘，而彈性網絡則會把相關的預測變數成組保留。

1.1.3 glmnet 套件

R 套件 glmnet 由史丹佛大學的 Jerome Friedman、Trevor Hastie 與 Robert Tibshirani 開發。它使用極其高效的座標下降演算法為廣義線性模型（generalized linear model, GLM）實作彈性網絡正則化（Friedman, Hastie, and Tibshirani 2010），其後又擴展到 Cox 比例風險模型（Simon et al. 2011）以及全部廣義線性模型族（Tay, Narasimhan, and Hastie 2023）。該套件可處理 gaussian（線性）、binomial（羅吉斯）、poisson 及其他 GLM 族。

glmnet 的主要特性包括：

- 正則化路徑：一次運算即可在整個 lambda 網格上擬合模型，比在每個 lambda 處分別擬合快得多。
- 交叉驗證（cross-validation）：`cv.glmnet()` 透過 k 摺交叉驗證自動選出最佳的正則化強度。
- 係數界限：支援對單個係數設定上限和下限（符號約束）。
- 鬆弛 lasso（relaxed lasso）：在選中的變數上以減弱的罰則或完全不加罰則重新擬合模型，以減小偏誤。

1.1.4 與 MARS (earth) 和 GAM (mgcv) 的比較

earth 應用程式及其兩個後處理模組（glmnet 與 mgcv）使用不同的建模引擎。下表總結了它們的主要差異：

特性	glmnet（彈性網絡）	earth (MARS)	mgcv (GAM)
模型類型	線性（帶正則化）	分段線性（自適應樣條）	平滑非線性（罰則樣條）
非線性	僅能透過交互項或基底函數展開實現	透過帶資料驅動節點的鉸鏈函數自動實現	透過平滑函數自動實現（薄板、三次樣條等）
變數選擇	內建，經由 L_1 罰則（Lasso）	內建，經由前向 / 後向逐步法配合 GCV 修剪	選用，經由雙重罰則收縮（ <code>select=TRUE</code> ）
交互項	兩兩交叉乘積（不易解釋）	基於鉸鏈函數的乘積（可解釋，最高 3 階）	張量積平滑項（ <code>te()</code> 、 <code>ti()</code> ）
可解釋性	係數即線性權重，逐個易於說明	g 函數按變數展示分段線性偏效應	平滑的偏效應曲線，非常直觀
最適用於	高維資料、變數選擇、小樣本	自動識別非線性與交互作用	平滑的非線性關係、靈活建模
RCA 調整	逐變數的線性調整	經由 g 函數的分段線性調整	經由偏效應的平滑調整
過擬合控制	交叉驗證選定的 lambda	基於 GCV 的修剪	REML/GCV 平滑參數估計

MARS (Multivariate Adaptive Regression Splines, 多元自適應迴歸樣條) 由 Jerome Friedman 於 1991 年提出。它透過從資料中自適應地選擇鉸鏈函數 $\max(0, x - k)$ 及其節點位置來建構分段線性模型。其 R 實作是 Stephen Milborrow 編寫的 **earth** 套件。MARS 的長處在於無需使用者指定即可自動發現非線性關係與交互作用。它的 **g** 函數 (按變數歸組的基底項) 提供了高度可解釋的偏效應曲線。

GAM (Generalized Additive Models, 廣義可加模型) 由 Hastie 與 Tibshirani 於 1990 年正式確立。其 R 實作 **mgcv** 由 Simon Wood 編寫, 提供帶自動平滑度選擇 (經由 REML 或 GCV) 的罰則迴歸樣條。GAM 產生的平滑偏效應曲線便於視覺化和解釋。當 **earthUI** 把節點位置匯出到 **mgcvUI** 時, **earth** 得出的節點會作為 GAM 平滑項的起始點, 從而把 MARS 的自適應節點定位與 GAM 的平滑估計結合起來。

1.1.5 建議的工作流程

對於不動產估價 (real estate appraisal) 以及其他需要可解釋、可辯護模型的類似應用：

1. 從 **earthUI** 入手, 發現資料中的重要變數、非線性關係與交互作用。**earth** 的自動基底函數選擇與 **g** 函數能提供極佳的初步洞察。
2. 用 **mgcvUI** 加以細化, 如果您希望得到更平滑的偏效應。把 **earthUI** 的節點匯入 **mgcvUI**, 即可從分段線性模型平順過渡到平滑模型。
3. 使用 **glmnetUI**, 當您需要力度更大的變數選擇 (Lasso)、預測變數數量接近或超過觀測數量, 或者希望得到一個帶正則化、便於辯護的純線性模型時。

1.2 什麼是 glmnetUI ?

glmnetUI 是 R 語言 **glmnet** 套件的圖形使用者介面 (graphical user interface)。它以本機 Shiny 應用程式的形式執行——無需登入、無需伺服器、也沒有帳號。您從 R 中啟動它, 匯入資料集 (CSV 或 Excel), 設定模型, 然後互動式地進行擬合。

從當前版本起, **glmnetUI** 不再是一個單獨安裝的套件：它作為彈性網絡後處理模組隨 **earthUI** 套件一起發布, 用 `earthUI::launch_glmnet()` 啟動。應用程式本身——以及本文所述的一切——並無改變；不同的只是安裝方式和啟動指令。三個常式 (**earth**、**glmnet**、**mgcv**) 共享同一個專案樹、同一個設定資料庫以及同一套支撐基礎設施。

該應用程式提供了完整的工作流程：資料匯入、變數設定、模型擬合、診斷圖、變數重要性、模型方程式, 以及可下載的 Word、PDF 或 HTML 格式報告。

彈性網絡迴歸結合了兩種正則化技術：

- 嶺迴歸 ($\alpha = 0$)：按比例把所有係數向零收縮。當您認為所有變數都相關時較為合適。
- Lasso 迴歸 ($\alpha = 1$)：可以把係數收縮到恰好為零, 從而實現自動變數選擇。產生的模型更簡潔、更易解釋。
- 彈性網絡 ($0 < \alpha < 1$)：兩種做法的混合。當預測變數之間相關時很有用。

1.3 三種用途模式

每個 **glmnetUI** 專案都有一個 **Purpose** (用途), 在建立專案時選定 (第 1 節, *Project* (專案))。當前專案的用途決定哪些工具和介面元素會出現；若要在另一種模式下工作, 請建立或開啟一個具有該用途的專案。三種模式為：

- **General** (通用) ——適用於任何類型的母體或資料集的彈性網絡迴歸。這是預設模式。它提供完整的正則化迴歸工作流程, 不含任何領域專用的附加功能。
- **For Appraisal** (用於估價) ——為不動產估價量身打造的彈性網絡迴歸。增加了針對單一不動產估值的功能, 包括估價標的處理、特殊欄位指定以及殘差約束法 (Residual Constraint Approach, RCA)。
- **Market Area Analysis** (市場區域分析) ——為市場區域研究量身打造的彈性網絡迴歸。增加了分析既定市場中成組不動產的功能, 並可選擇排除標的列。

在這三種模式下，核心建模引擎完全相同——您始終是在擬合一個彈性網絡（glmnet）模型。用途設定只控制哪些附加工具和介面元素可用。

1.4 不動產專用功能

當選擇 **For Appraisal** 或 **Market Area Analysis** 時，glmnetUI 會啟用若干為不動產分析設計的功能：

- 特殊欄位指定——每個預測變數都可以標記一個特殊角色，例如 `contract_date`、`dom`、`concessions`、`latitude`、`longitude`、`living_area`、`lot_size`、`actual_age`、`effective_age`、`area`、`site_dimensions` 或 `display_only`。這些指定控制該欄在擬合與輸出過程中的處理方式。
- 經緯度取整——被指定為緯度或經度的欄會自動四捨五入到小數點後 3 位，以防止過擬合。
- **Sale Age** 欄（成交時距）——當某欄被指定為 `contract_date` 且提供了 `Effective Date`（估價基準日）時，glmnetUI 會計算一個 `sale_age` 欄（成交日期與基準日之間的天數），並用它替代原欄作為預測變數。
- **RCA** 計算（僅限估價模式）——在估價模式下，擬合模型之後 glmnetUI 可以計算殘差約束法（RCA）輸出，得出逐個可比案例的調整值、淨 / 毛調整彙總，以及估價標的的調整後成交價格。

提示：*General* 用途模式適用於任何資料集——金融、科學、工程或不動產。估價模式與市場模式只是為不動產專業人士增加了一些便利功能。

1.5 快速上手

使用 glmnetUI 的步驟：

1. 在 R 中安裝 **earthUI**：`install.packages("earthUI")` ——glmnet 常式已包含在內。（關於開發版和選用元件，參見本期 earthUI 一文中的「Installing earthUI and vProlog」。）
2. 啟動應用程式：在 R 中執行 `earthUI::launch_glmnet()`。應用程式會在您的網頁瀏覽器中以連接埠 7879 開啟。您也可以直接在瀏覽器中前往 <http://localhost:7879>。應用程式會記住您上次使用的用途模式並自動還原。
3. 在側邊欄第 1 節建立或開啟專案。專案設定用途和工作地點（國家、州 / 省、縣、市）及其輸入 / 輸出資料夾。專案存放在共享的 `regProj` 根目錄下，並與姊妹應用程式 earthUI 和 mgcvUI 共用。開啟一個具有不同用途的專案，就是切換模式的方式。
4. 在第 2 節匯入資料，從當前專案的輸入資料夾中選擇一個 CSV 或 Excel 檔案。把檔案放入該資料夾後，點選 **Refresh**（重新整理）即可列出它們。
5. 從 **earthUI** 匯入（選用）——第 3 節允許您匯入一個 earthUI 結果 `.rds` 檔案，以便把 earth 的鉸鏈基底函數與 glmnet 的正則化結合使用。若沒有 earthUI 結果，可跳過此步。
6. 設定變數——選擇目標變數和預測變數，設定資料型別、預期符號，並指定任何特殊欄位角色。
7. 設定 **glmnet** 參數——`alpha`、`lambda`、`family`、符號約束、交互項及其他選項。
8. 擬合模型——點選「Fit Glmnet Model」（擬合 Glmnet 模型），並在主面板中檢視結果。
9. 匯出——將預測結果下載為 Excel，產生並轉換 Quarto 報告，或（在估價模式下）計算 RCA 調整值與銷售比較表格（Sales Comparison Grid）。

設定會自動儲存在瀏覽器的本機儲存空間中，並在您重新載入同一個輸入檔案時還原。

2 MLS 輸入資料要求

對於不動產估價與市場分析工作流程，您的輸入資料通常來自多重上市服務（Multiple Listing Service，MLS）匯出檔案。本章說明預期的檔案結構以及 glmnetUI 可以使用的欄位。

2.1 檔案格式與結構

glmnetUI 接受 CSV 與 Excel (.xlsx、.xls) 檔案。匯入時，欄名會自動轉換為 snake_case——例如，「Living SqFt」變為 living_sqft，「Contract Date」變為 contract_date，「Sale Price」變為 sale_price。這種規範化確保了整個工作流程中欄位參照的一致性。匯入時使用的 CSV 分隔符號和小數點記號由地區設定決定（參見第 3 章「Locale & Regional Settings」）。

您的資料檔案應當是一張平面表格，每列一筆不動產，每欄一個屬性。檔案的第一列必須包含欄標題。

2.2 估價模式所需的欄位

雖然 glmnetUI 可以處理任意一組欄位，但完整的估價工作流程若具備以下各欄會更為得力：

欄位	特殊型別	用途
Sale Price（成交價格）	（目標變數）	模型的反應變數
Contract Date（合約日期）	contract_date	用於計算 sale_age（距基準日的天數）
Listing Date（掛牌日期）	listing_date	在沒有 DOM 欄時，與合約日期一起計算 DOM
Days on Market（在市天數）	dom	在市天數；顯示於匯出結果中
Concessions（成交讓利）	concessions	成交讓利；淨成交價 = 成交價格 - 讓利
Living Area (SF)（居住面積）	living_area	啟用每平方英尺殘差 (residual_sf、cqa_sf)
Lot Size（基地面積）	lot_size	基地面積欄
Site Dimensions（基地尺寸）	site_dimensions	與基地面積歸為一組
Latitude（緯度）	latitude	四捨五入到小數點後 3 位
Longitude（經度）	longitude	四捨五入到小數點後 3 位
Area ID（區域編號）	area	市場區域 / 街廓識別碼
Actual Age（實際屋齡）	actual_age	不動產的屋齡
Effective Age（有效屋齡）	effective_age	不動產的有效屋齡
Address（地址）	display_only	顯示於匯出結果中；不進入模型

試算表的欄名可以使用外語——「特殊」名稱之所以用英文，是為了讓 R 程式能對它們作特殊處理。除此之外，您所給定的欄名會出現在迴歸模型、圖形中，以及（若在做估價）輸出報告中。

並非所有欄位都是必需的。glmnetUI 會自行調適——若缺少某一欄，相應的功能就直接省略。不過，對於不動產定價模型，為達到可接受的擬合效果，強烈建議具備以下幾欄：

1. **Sale Age** (成交時距) ——合約成交日期與估價或分析基準日之間的天數。若使用跨年度的成交歷史資料，尤其是超過 5 年的期間，`sale_age` 在估計成交價格時往往起核心作用。
2. **Living Area** (居住面積) ——也常被稱為「Living Sqft」、「GLA」(gross living area，總居住面積) 等。這是成交價格的又一主要決定因素。
3. **Total Bath Count** (衛浴總數) ——全套、四分之一套、半套和四分之三套衛浴的總數。例如，兩套全衛浴加一套半衛浴記為 2.5。
4. **Garage Bays** (車位數) 或 **Garage Area** (車庫面積) ——車庫停車位數量或車庫面積。
5. **Lot Size** (基地面積) ——不動產的土地面積，通常以平方英尺或英畝計。
6. **Longitude** (經度)、**Latitude** (緯度)，以及 (若有) **Area ID** (區域編號)。這些位置變數有助於模型解釋地理上的價格差異。

2.3 特殊欄位命名慣例

glmnetUI 透過特殊型別指定而非欄名來識別欄位。在 MLS 匯出檔案中，您可以隨意命名各欄——關鍵在於您在變數設定表 (第 6 章) 中指定了正確的特殊型別。

例如，您的 MLS 可能把居住面積匯出為「GLA」、「Living SqFt」、「liv_area」或「gross_living_area」。匯入之後 (欄名轉為 snake_case)，您只需在 Special (特殊) 下拉式選單中把它指定為 `living_area`。此後無論其原始名稱如何，glmnetUI 都會用它來計算每平方英尺的殘差。

2.4 資料品質與完整性

- 遺漏值 (NA)：任何預測變數欄或目標欄中含 NA 的資料列，會在擬合前自動剔除。
- 日期欄：符合常見日期格式的字元欄在匯入時會自動解析。當偵測不到 4 位年份時，優先按 2 位年份格式解析。
- 數值欄：成交價格、居住面積、基地面積、讓利等欄位必須為數值。若您的 MLS 匯出的價格帶有貨幣符號或千位分隔符號 (例如「\$350,000」或「350.000,00」)，可能需要在匯入前先行清理。
- 因子欄：文字欄、TRUE/FALSE 欄和 R 因子欄會自動按類別變數處理——迴歸只能把它們當作因子。而數值欄 (例如數字形式的區域碼或風格碼) 只有在您勾選其 **Factor** (因子) 核取方塊時才會變成因子；glmnetUI 絕不會根據數值的取值範圍推斷其為類別變數，因此像 `bath_count` 這樣的離散數值預測變數仍保持連續。Special 角色與是否作為因子無關。

提示：匯入後請檢視預測變數表中的 NA 一欄。遺漏值很多的欄位可能導致大量資料列被剔除。可以考慮排除高 NA 欄位，或在匯入前清理資料。

2.5 標的列的位置

在 **Appraisal** (估價) 模式下，第 1 列必須是估價標的。其餘各列均為可比成交案例。標的列不參與模型擬合。擬合完成後，模型仍會為標的列產生預測值。

在 **Market Area Analysis** (市場區域分析) 模式下，把標的放在第 1 列是選用的。

在 **General** (通用) 模式下沒有特殊的資料列處理——所有資料列一視同仁。

3 通用模式

3.1 概述

啟動 glmnetUI 時的預設模式即為通用 (General Purpose) 模式。它為任何資料集——而不只是不動產資料——提供完整的彈性網絡迴歸工作流程。您可以把 glmnetUI 用於科學資料、金融分析、工程研究，或任何迴歸問題。

在通用模式下，介面會略去不動產專用功能（特殊欄位、成交時距、座標取整、RCA）。側邊欄經過精簡，專注於變數選擇、參數設定、模型擬合與匯出。

3.1.1 跳過第一列

開啟通用或市場專案後，側邊欄頂部會出現一個 **Skip first row**（跳過第一列）核取方塊。勾選後，第 1 列將不參與模型擬合。當第 1 列包含一個不應影響模型的標的或參照觀測值時，這一選項很有用。在估價模式下，第 1 列（估價標的）始終自動排除。

3.1.2 按用途分別儲存設定

設定（預測變數選擇、模型參數、交互項）會按輸入檔案與用途模式的每種組合分別儲存。在通用、估價和市場模式之間切換時，各模式的設定互不影響、各自保留。

3.2 側邊欄工作流程

側邊欄按編號、可收合的小節組織，引導您從專案選擇一路走到報告匯出。多數小節只在開啟使用中的專案後才出現，其中幾個還與專案的用途有關。輸出資料夾由使用中的專案推導得出——沒有單獨的輸出資料夾欄位。

1. **Project**（專案）——建立或開啟專案。專案設定用途和工作地點（國家、州 / 省、縣、市）及其輸入（<os>_in/）和輸出資料夾。專案存放在共享的 **regProj** 根目錄下（預設 `~/regProj`，可用 `REGPROJ_ROOT` 環境變數覆寫），並與姊妹應用程式 **earthUI** 和 **mgcvUI** 共用，因此三者的模型可以並存。點選 **Close Project**（關閉專案）即可切換。

2. **Import Data**（匯入資料）——從使用中專案的輸入資料夾中選擇一個 CSV 或 Excel 檔案。透過 Finder / 檔案總管把檔案放入該資料夾，然後點選 **Refresh**（重新整理）列出它們。對於含多個工作表的 Excel 檔案，會出現工作表選擇器。欄名會自動轉換為 `snake_case`。

3. **Import from earthUI**（從 earthUI 匯入，選用）——匯入一個 earthUI 結果 `.rds` 檔案，以便把 earth 的鉸鏈基底函數與 glmnet 的彈性網絡正則化結合使用。瀏覽按鈕的樣式與第 2 節的檔案選擇器一致。此步驟為選用——若沒有可匯入的 earthUI 結果，可以跳過。

4. **Variable Configuration**（變數設定）——目標變數選擇器，帶 Include（納入）、Factor（因子）和 Force（強制保留）核取方塊的預測變數表，以及 Sign（符號）下拉式選單。Special（特殊）欄——用於指定合約日期、座標等——只在估價與市場模式下出現。詳見第 6 章。

5. **glmnet Call Parameters**（glmnet 呼叫參數）——全部模型設定：alpha、lambda、family、standardize、符號約束、鬆弛 lasso、隨機種子、交互項矩陣，以及進階參數（lambda.min.ratio、nlambdas、CV 損失度量、收斂閾值、最大疊代次數、截距）。完整的參數參考見第 7 章。

6. **Fit Glmnet Model**（擬合 Glmnet 模型）——執行模型的按鈕。結果會顯示在右側的各個分頁中。

7. **Download Output**（下載輸出）——把預測值、殘差、CQA 分數和逐變數貢獻匯出為 Excel 檔案，存入專案的輸出資料夾。模型擬合後，在所有用途模式下均可用。

8. **Calculate RCA Adjustments & Download**（計算 RCA 調整值並下載）——為每個可比案例計算逐變數的 RCA 調整值，並透過 CQA 內插得出標的殘差。僅在估價與市場模式下、擬合之後出現。詳見第 11 章。

9. Generate Sales Grid & Download (產生銷售比較表格並下載) ——產生中間的銷售比較表格，按毛調整百分比對可比案例排序。僅在估價模式下、計算完 RCA 調整值之後出現。詳見第 12 章。

10. Generate Quarto Report (產生 Quarto 報告) ——把一個自足的 Quarto (.qmd) 報告包——含全部圖形與表格——寫入專案的輸出資料夾。模型擬合後即可使用。

11. Convert Quarto Report (轉換 Quarto 報告) ——把產生的 .qmd 報告包轉譯為 HTML、Word 或 PDF。PDF 需要安裝 LaTeX；若未偵測到，則隱藏 PDF 選項。詳見第 13 章。

在市場模式與通用模式下，僅限估價的步驟會被跳過，兩個報告小節的編號也相應調整。

3.3 專案

glmnetUI 把工作組織為共享 regProj 根資料夾下的專案 (earthUI 與 mgcvUI 使用的是同一棵目錄樹)。每個專案擁有一個輸入資料夾 (<os>_in) 和按引擎劃分的輸出資料夾 (例如 <os>_out_glmnet)。

要建立專案，請點選第 1 節中的 + New... (新建)，並提供：

- **Purpose** (用途) ——General、Appraisal 或 Market Area。
- **Location** (地點) ——專案的政治 / 行政層級。需要填寫哪些層級取決於您選擇的國家，因此各國的命名慣例都能自動支援——美國專案使用州 / 縣 / 市，法國專案使用 region/département/commune，日本專案使用都道府縣 / 市或區 / 町丁，依此類推。
- **Project Name** (專案名稱) ——一個簡短標籤，限 8 個字元 (字母、數字、_ 和 -)。

應用程式不會原樣使用您輸入的名稱：它會在前面加上地點代碼和建立日期時間，構造出一個唯一的帶時間戳記的資料夾名，形式為 <country>_<admin level codes>_<YYYYMMDD-HHMMSS>_<your name> ——例如，一個名為 burl 的美國專案會變成 us_ca_081_burlin_20260604-131500_burl。8 個字元的限制只適用於您鍵入的那一部分。只有新建的專案才會帶上時間戳記前綴；磁碟上已有的專案保留原名，並且照常完全可用。

3.4 從 earthUI 匯入 (選用)

側邊欄的第 3 節——**Import from earthUI** (從 earthUI 匯入) ——允許您匯入一個 earthUI 結果 .rds 檔案。這使得 glmnetUI 能夠把 earth 的鉸鏈基底函數 (分段線性項) 與 glmnet 的彈性網絡正則化結合起來，把 earth 的自動非線性識別與 glmnet 的變數選擇和係數約束合二為一。

匯入 earthUI 結果後，驅動 glmnet 擬合的是 earth 模型的預測變數，而不是 Include 核取方塊，因為 glmnet 是在 earth 的基底函數展開上擬合的。變數設定表會明確體現這一點：**Include** 與 **Type** 被鎖定，earth 模型未使用的預測變數所在資料列被停用，並有一條提示說明預測變數集來自 earth 模型。**Force** 與 **Sign** 對 earth 的預測變數仍可編輯，因為它們在擬合時仍會作用於 earth 基底函數。

瀏覽按鈕的樣式與第 2 節的檔案選擇器一致。此步驟完全選用——若沒有 earthUI 結果檔案，直接跳過第 3 節進入第 4 節即可。

3.5 主面板分頁

擬合完成後，主面板提供以下分頁：

分頁	內容
Data Preview (資料預覽)	匯入資料的預覽。在估價模式下，分為估價標的表與可比成交案例表。
Equation (方程式)	以 LaTeX/MathJax 顯示的擬合模型方程式，展示每個非零係數與交互項。
Correlation (相關性)	數值型預測變數相關係數的熱圖（擬合前即可檢視）。
Summary (摘要)	關鍵指標 (R^2 、調整 R^2 、 GR^2 、 $CV R^2$ 、RMSE、MAE) 與估價指標 (COD、PRD、中位數比率)。
Coefficients (係數)	非零係數表，附符號違例警示。
Variable Importance (變數重要性)	預測變數重要性 ($ \beta \times sd(x)$) 的長條圖與排序表。可用時支援 plotly 互動式滑鼠停留。
Contributions (貢獻)	帶斜率標註的逐變數偏效應圖。數值型預測變數為散布圖 + 擬合線，因子為盒鬚圖，交互項為三維曲面 / 散布圖 / 熱圖。
ANOVA (變異數分析)	變異數分解表：逐變數的平方和、估模型平方和的百分比、係數。
Diagnostics (診斷)	四幅圖：係數路徑、CV 誤差、實際值對預測值、殘差對擬合值。
Glmnet Output (Glmnet 輸出)	原始模型輸出：所用隨機種子、模型打印輸出、選定的 λ 、 $\lambda_{min}/1se$ 、 γ (若使用鬆弛 lasso)，以及完整的係數向量。
Report (報告)	報告設定與匯出欄位。
RCA Adjustments (RCA 調整)	RCA 分析結果（估價模式，計算完成後）。

3.6 設定的保存

glmnetUI 會自動把您的設定儲存到瀏覽器的本機儲存空間中，以輸入檔名作為索引鍵。當您重新載入同一個檔案時，所有設定都會還原：目標變數選擇、預測變數核取方塊、資料型別、預期符號、glmnet 參數以及交互項矩陣。上次使用的用途模式也會全域儲存，並在重新啟動應用程式時還原。有三個選項可供選擇：

- **Use last settings for input file** (使用該輸入檔案的上次設定) —— 還原為此檔案儲存的設定
- **Use default settings** (使用預設設定) —— 套用您儲存的全域預設值
- **glmnet defaults** (glmnet 預設值) —— 重設為出廠預設值 ($\alpha = 1$ 、 $CV \lambda_{1se}$ 、10 摺、gaussian、開啟標準化)

Save current as default (將目前設定存為預設) 按鈕可把目前的全部設定儲存為全域預設值。

3.7 深色模式

點選右上角的月亮 / 太陽圖示，即可在淺色與深色主題之間切換。主題偏好儲存在本機儲存空間中，跨工作階段保持。所有介面元素（表格、圖形、卡片、按鈕）都會透過 CSS 變數適配所選主題。

3.8 地區與區域設定

glmnetUI 透過基於國家的地區設定系統支援各國的數字與 CSV 格式慣例。標題列中的 **Settings** (設定) 下拉式選單提供面向 31 個支援國家的 **Country** (國家) 與 **Paper** (紙張)

選擇器。每個預設會設定：

- **CSV 分隔符號**——美國 / 英國 / 日本用逗號 (,)，歐洲多數國家用分號 (;)，因為那裡逗號被用作小數點。
- **小數點記號**——句點 (.) 或逗號 (,)。
- **千位分隔符號**——逗號 (美國 / 英國 / 日本)、句點 (德國 / 義大利 / 西班牙)、空格 (芬蘭 / 法國 / 波蘭 / 波羅的海國家 / 烏克蘭 / 俄羅斯)，或撇號 (瑞士)。
- **紙張尺寸**——Letter (美國 / 加拿大 / 墨西哥) 或 A4 (其他所有地區)。

3.8.1 儲存預設值

點選 **Save as my default** (存為我的預設值)，即可全域儲存您的地區設定偏好。這些預設值將適用於今後所有工作階段，無論您載入哪個資料檔案。

提示：圖形座標軸和斜率標註上的數字格式會自動適配您的地區設定。德語地區用句點作千位分隔符號 (200.000)，芬蘭用空格 (200 000)，瑞士用撇號 (200'000)。不顯示任何貨幣符號——*glmnetUI* 與幣別無關。

4 估價模式

當您把 Purpose 選為 **For Appraisal**（用於估價）時，glmnetUI 會針對單一不動產估值進行自我設定。第 3 章所述的全部功能依然可用；本章只講估價模式特有的附加內容。

4.1 標的列的處理

在估價模式下，資料集的第 1 列即估價標的，其餘各列均為可比成交案例。您的輸入檔案必須據此組織（參見第 2 章）。標的的成交價格可以留空或填任意值——glmnetUI 在擬合時會自動把它當作 NA 處理。

匯入之後，Data Preview 分頁會分為兩部分：**Subject Property**（估價標的，第 1 列）與 **Comparable Sales**（可比成交案例，第 2 列及以後）。第 1 列始終不參與模型擬合。擬合完成後，模型仍會為標的列產生預測值。

4.2 基準日與成交時距

在估價與市場模式下，變數設定小節中會出現一個 **Effective Date**（估價基準日）欄位（預設為當天日期）。如果您在 Special 欄下拉式選單中把某欄指定為 `contract_date`，glmnetUI 就會計算一個 `sale_age` 欄——即每筆成交的合約日期與基準日之間的整數天數。該欄會作為預測變數加入模型。

當基準日發生變化時，`sale_age` 會自動重新計算。

4.3 特殊欄位指定

在估價與市場模式下，變數設定表中每個預測變數都會出現一個 **Special**（特殊）下拉式選單。特殊型別及其作用的完整參考見第 6 章。

4.4 RCA 調整概覽

Calculate RCA Adjustments & Download（計算 RCA 調整值並下載）按鈕（側邊欄第 8 節，在估價與市場模式下擬合之後可見）會為每個可比案例計算相對於標的的市場推導調整值。完整的 RCA 工作流程見第 11 章。

提示：您賦予標的的 *CQA* 分數決定了殘差分布中有多大比例歸屬於標的。分數為 5.00 時，標的位於各可比案例的中位數處。

5 市場區域分析模式

當您把 Purpose 選為 **Market Area Analysis**（市場區域分析）時，glmnetUI 提供與估價模式相同的不動產專用功能（特殊欄位、成交時距、座標取整），但其取向是分析一組不動產，而非為單一標的估值。

5.1 與估價模式的差異

- 標的列處理更為靈活——在市場模式下，第 1 列不會被自動當作估價標的。
- 沒有銷售比較表格——「Calculate RCA Adjustments & Download」（第 8 節）步驟在市場模式下可用，但銷售比較表格步驟（第 9 節）僅限估價模式。

5.2 何時使用市場模式

在以下情形下適合採用市場區域分析模式：

- 為某個街廓或市場區域建立迴歸模型，以了解價值驅動因素
- 分析面積、屋齡、基地面積和位置等變數如何影響一組不動產的成交價格
- 為市場狀況分析或街廓界定準備佐證資料
- 處理某個資料集時，您需要特殊欄位功能（成交時距、座標取整），但不需要 RCA 調整工作流程

提示：當您需要特殊欄位功能但不需要 RCA 調整工作流程時，市場模式同樣適用於一般性的不動產迴歸分析。

6 變數選擇

側邊欄的第 4 節——**Variable Configuration**（變數設定）——是您選擇哪些欄位參與模型以及如何處理它們的地方。

6.1 目標變數

第 4 節頂部的 **Target (response) variable**（目標 / 反應變數）下拉式選單會列出資料集中的每一欄。請選擇一欄作為反應變數（例如成交價格）。目標欄會自動從預測變數清單中排除。

6.2 預測變數表

在目標變數選擇器下方，有一張表列出其餘每一欄，包含以下欄位：

欄位	說明
Variable （變數）	欄名
Type （型別）	資料型別下拉式選單：numeric、integer、character、Date、POSIXct
Include （納入）	核取方塊——把該欄作為預測變數納入模型
Factor （因子）	核取方塊——把該變數當作類別因子處理（在模型矩陣中產生虛擬變數欄）
Force （強制保留）	核取方塊——強制進入模型（罰則因子 = 0，絕不會被 lasso 剔除）
Special （特殊）	下拉式選單（僅估價 / 市場模式）——參見下文的特殊欄位型別參考
Sign （符號）	預期的係數符號：正、負或不限
NAs	遺漏值計數

6.3 資料型別偵測與手動覆寫

glmnetUI 在匯入時會自動偵測資料型別，能夠識別數值型、整數型和日期型的欄位。看起來像日期的字元欄（符合常見日期格式模式）會被歸類為 **Date**。

您可以透過修改 **Type** 下拉式選單來覆寫任何偵測結果。更改型別會影響該欄在模型矩陣中的編碼方式。

文字欄、TRUE/FALSE 欄和 R 因子欄會自動按類別變數處理。數值欄只有在您勾選其 **Factor** 核取方塊時才會變成因子——glmnetUI 絕不會根據數值的取值範圍推斷其為類別變數，因此像衛浴數量這樣的離散數值預測變數除非您另行指定，否則保持連續。Factor 設定透過 localStorage 持續保存。

6.4 預期符號與約束的施加

對每個預測變數，**Sign** 下拉式選單用於指定您預期其係數為正、為負還是無限方向。擬合之後：

- 符號與預期不符的係數會在係數表中被標記出來
- 符號警示會顯示在 Summary 分頁中

當 glmnet 參數（第 7 章）中啟用了 **Enforce Sign Constraints**（施加符號約束）核取方塊時，模型會被強制產生與預期符號一致的係數。這是透過 glmnet 中的 upper.limits 與 lower.limits 實現的。

6.5 特殊欄位型別參考

在估價與市場模式下，**Special** 下拉式選單提供以下選項：

Weighting (加權)：

- **weight** ——觀測權重欄（所有模式下均可用；只允許指定一欄；權重為 0 的資料列不參與擬合）

Date & Time Types (日期與時間型別)：

- **contract_date** ——觸發依據基準日自動計算 **sale_age**
- **listing_date** ——在計算在市天數時作為備用來源
- **dom** ——標識在市天數 (Days on Market) 欄
- **sale_age** ——指定一個已有的成交時距欄，這樣任何一欄都可以充當成交時距，而不必讓 glmnetUI 從 **contract_date** 計算

Monetary Types (貨幣型別)：

- **concessions** ——標識成交讓利（賣方補貼、買方誘因等）

Transaction Types (交易型別)：

- **sale_type** ——標識成交 / 交易型別欄

Size & Location Types (面積與位置型別)：

- **latitude** ——數值自動四捨五入到小數點後 3 位
- **longitude** ——與緯度作相同的取整處理
- **area** ——市場區域或街廓識別碼（例如 **area_id**、**neighborhood**、**market_area**）
- **living_area** ——啟用按平方英尺計的殘差計算（**residual_sf** 與 **cqa_sf**）
- **lot_size** ——基地面積欄
- **site_dimensions** ——與基地面積歸為一組（例如「75x120」）

Age Types (屋齡型別)：

- **actual_age** ——不動產的屋齡欄
- **effective_age** ——不動產的有效屋齡

Display Types (僅顯示型別)：

- **display_only** ——該欄會包含在 Excel 匯出結果中，但完全不參與模型擬合。可用於地址欄位、MLS 編號、地號或其他參考資料。可以有多欄採用此指定。

提示：被指定為 *display_only* 的欄位仍保留在資料集中並出現在 Excel 匯出結果裡，但完全不進入預測變數表。可用於 ID 欄、地址或其他參考欄位。

7 參數選擇

側邊欄的第 5 節——**glmnet Call Parameters** (glmnet 呼叫參數) ——提供彈性網絡模型的全部設定選項。每個參數旁都有一個藍色說明圖示 (?)，滑鼠停留可看到說明。

7.1 Alpha (混合參數)

alpha 參數控制正則化的類型：

Alpha 取值	類型	行為
0	嶺迴歸	按比例收縮所有係數；絕不會把任何係數置為零
1	Lasso	可以把係數置為恰好為零 (變數選擇)
0 至 1	彈性網絡	嶺迴歸與 lasso 的混合

有兩種 alpha 選擇方法：

- **Fixed** (固定) ——透過滑桿設定單一 alpha 值 (預設：1.0)
- **Grid Search** (網格搜尋) ——測試多個 alpha 值，選出 CV 誤差最小的那個。可設定取值範圍與取值個數。

7.2 Lambda 的選擇

lambda 控制正則化的強度：

方法	說明
交叉驗證	推薦。使用 k 摺 CV 測試多個 lambda 值。
手動	若您有充分理由，可輸入指定的 lambda 值。

使用交叉驗證時，有兩種 lambda 取法：

- **lambda.1se** (預設) ——在最小值一個標準誤範圍內的最大 lambda。產生更簡潔、正則化更強的模型。估價工作推薦使用。
- **lambda.min** ——使交叉驗證誤差最小的 lambda。產生擬合最好但可能更複雜的模型。

Number of CV Folds (CV 摺數) 參數控制交叉驗證時資料的劃分方式 (預設：10 摺)。

7.3 Family (分布族)

為您的反應變數選擇分布族：

分布族	適用情形
gaussian	連續型反應 (例如成交價格)。最常用。
binomial	二元結果 (例如已售 / 未售)。
poisson	計數資料 (例如成交筆數)。

7.4 Standardize (標準化)

勾選時 (預設)，所有預測變數在擬合前都會被縮放為平均數 0、標準差 1。這確保罰則對所有預測變數一視同仁，而不受其原始尺度影響。係數會按原始尺度傳回。通常應保持開啟。

7.5 符號約束

啟用 **Enforce Sign Constraints** 後，係數會被約束為與變數表（第 6 章）中設定的預期符號一致：

- 「positive」（正）符號強制係數 ≥ 0
- 「negative」（負）符號強制係數 ≤ 0
- 設為「either」（不限）的變數不受約束

這是透過 `glmnet()` 中的 `upper.limits` 與 `lower.limits` 參數實現的。

7.6 鬆弛 Lasso

常規 lasso 對變數選擇和係數估計使用同一個罰則，這可能把重要係數過度收縮到零。鬆弛 lasso (relaxed lasso) 把這兩步分開：

1. 首先，透過 lasso 選擇變數（由罰則決定哪些係數得以保留）
2. 然後，對保留下來的變數以更小的罰則或不加罰則重新擬合，得到偏誤更小的估計

Gamma 滑桿控制鬆弛的程度：

Gamma	效果
0	完全鬆弛（OLS 重新擬合，完全不作收縮）
1	不作鬆弛（與常規 glmnet 相同）
介於其間	在鬆弛擬合與帶罰則擬合之間取折衷

在使用交叉驗證時，gamma 會被自動選取以獲得最佳表現。

7.7 隨機種子

Random seed（隨機種子）文字方塊預填了一個隨機整數，用於使交叉驗證的摺分配可重現——擬合前會呼叫 `set.seed()`，所用的種子會顯示在 `Glmnet Output` 分頁以及擬合狀態訊息中。每次擬合之後會自動產生一個新種子（使用系統時間，而非 R 的亂數產生器）。目前檔案最近使用的 5 個種子會以可點選連結的形式列出，便於快速調回；載入新檔案時種子歷史會重設。

7.8 交互項矩陣

一個可互動的上三角矩陣讓您控制哪些變數對之間允許產生交互作用。每個儲存格內有一個核取方塊——勾選表示交互項 `x1:x2` 會被納入模型矩陣。

- **Allow All**（全部允許）——勾選所有交互變數對
- **Clear All**（全部清除）——取消勾選所有變數對（新檔案的預設狀態是全部不勾選）
- **點選變數名**——切換該變數的所有交互項

交互項的選擇會按輸入檔案分別儲存到瀏覽器 `localStorage`，並在您重新載入同一檔案時還原。

只要啟用了任何交互項，匯出的報告中就會包含一張 **Enabled Interactions**（已啟用交互項）矩陣（兩個座標軸都是預測變數，納入 `x1:x2` 乘積項之處打勾），以橫向頁面呈現，與 `earthUI` 的 `Allowed Interactions` 矩陣相對應。請注意，lasso 仍可能把某個已啟用項的係數收縮為零。

提示：當預測變數很多時，lasso 會自動把不需要的交互項歸零。先用「*Allow All*」起步，再讓模型自行選擇，是一種合理的做法。

警告

啟用 glmnet 交互項會產生在法庭或稽核場合難以解釋和說明的調整值。請謹慎使用交互項，並核實所得調整值是否合理。

7.9 禁止進入主效應（僅作交互項）

在交互項矩陣中對某個變數名按右鍵，即可禁止它以主效應的身份進入模型。變數名後會出現一個粗體的「1」標記。被禁止時：

- 該變數會從主公式項中移除
- 該變數仍可出現在交互項（x1:x2）中
- 再次按右鍵即可解除禁止
- 該狀態按輸入檔案透過 localStorage 持續保存

一個典型用例：在為應隨不動產面積按比例變化的時間調整建模時，禁止 sale_age 進入主效應，只允許它與 living_area 產生交互。這樣得到的是 sale_age:living_area 交互項，而不是一個平坦的 sale_age 係數——即時間增值率隨房屋面積而異。

7.10 進階參數

glmnet 呼叫參數底部有一個可收合的「Advanced」（進階）小節，其中是一些附加設定。它們採用合理的預設值，但仍然可見，以便所有模型設定都能為法庭或稽核留下書面記錄。

- **Lambda Min Ratio** —— 正則化路徑中最小 lambda 與最大 lambda 之比。預設值：0.00001。對於小資料集（觀測數 <50），取較大的值（例如 0.01）可能表現更好。
- **Number of Lambda Values** —— 路徑中 lambda 值的個數。預設值：100。
- **CV Loss Metric** —— 交叉驗證所用的損失函數：MSE（gaussian 的預設值）、MAE（對離群值穩健）或 Deviance（偏差）。
- **Convergence Threshold** —— 座標下降的收斂閾值。預設值：1e-07。
- **Max Iterations** —— 座標下降的最大疊代次數。預設值：100,000。若出現收斂警告，可調大此值。
- **Fit Intercept** —— 是否包含截距（常數）項。預設：開啟。在估價 RCA 中，截距即「基準值」。

7.11 設定預設值

有三個選項控制載入檔案時參數如何初始化：

- **Use last settings for input file** —— 還原上次處理此檔案時所用的確切設定
- **Use default settings** —— 套用您儲存的全域預設值
- **glmnet defaults** —— 重設為出廠設定（alpha = 1、CV lambda.1se、10 摺、gaussian、開啟標準化）

點選 **Save current as default**，即可把目前全部設定存為今後檔案所用的全域預設值。

8 擬合模型

8.1 擬合按鈕

側邊欄的第 6 節包含 **Fit Glmnet Model** (擬合 Glmnet 模型) 按鈕。點選它即以您目前的設定執行模型。

8.2 擬合過程中發生了什麼

當您點選 Fit 時，glmnetUI 會：

1. 準備資料——剔除含 NA 的資料列，施加權重 (若已指定)，排除標的列 (估價模式下)，並把因子變數編碼為虛擬變數欄。
2. 建構模型矩陣——產生設計矩陣，含交互項矩陣中所有被允許的交互項。
3. 施加約束——若啟用了符號約束，則為係數設定 `upper.limits` 與 `lower.limits`。
4. 執行交叉驗證——若選擇了 CV，則呼叫 `cv.glmnet()` 尋找最佳 `lambda`。若啟用了鬆弛 lasso，則呼叫 `cv.glmnet(..., relax = TRUE)`。
5. 擬合最終模型——儲存擬合好的模型、係數、預測值與殘差。
6. 更新所有結果分頁——填入 Summary、Coefficients、Equation、Variable Importance、Contributions、ANOVA 與 Diagnostics。

Fit 按鈕下方的狀態訊息會顯示觀測數量、被排除的資料列數以及選定的 `lambda`。

9 結果分頁

9.1 Data Preview (資料預覽)

以互動式 DataTable 顯示匯入的資料。在估價 / 市場模式下，預覽分為兩張表：**Subject Property** (估價標的，第 1 列) 與 **Comparable Sales** (可比成交案例，第 2 列及以後)。

每個儲存格都保持單行顯示，並按欄寬截斷，因此較寬的自由文字欄位 (如不動產備註) 不會把資料列撐得很高。按兩下任一儲存格可彈出視窗，顯示其完整、未截斷的內容。

9.2 Equation (方程式)

顯示透過 MathJax 以 LaTeX 轉譯的擬合模型方程式。展示所有非零係數，並採用規範的數學排版：

- 正係數以 + 號顯示
- 負係數以 - 號顯示
- 交互項顯示為 $x_1 \times x_2$
- 數字帶千位逗號，並保留適當的小數位數

9.3 Correlation (相關性)

一張熱圖矩陣，展示所有數值型預測變數與反應變數之間的 Pearson 相關係數。資料匯入後立即可用——無需擬合模型。

- 色彩漸層：藍色 (-1) → 白色 (0) → 紅色 (+1)
- 每個儲存格中印出相關係數值
- 文字與座標軸尺寸隨變數數量自適應
- 有助於在擬合前識別多重共線性

9.4 Summary (摘要)

以卡片形式顯示的模型擬合統計量：

- R^2 —— 被解釋變異數的比例
- $\text{Adj } R^2$ —— 按預測變數個數調整後的 R^2
- GR^2 —— 廣義 R^2 ，由 glmnet 在選定 lambda 處的偏差比得出
- $\text{CV } R^2$ —— 交叉驗證 R^2 (使用 CV 時)
- RMSE —— 均方根誤差
- MAE —— 平均絕對誤差

在估價 / 市場模式下，還會顯示以下卡片：

- **Median Ratio** —— (預測值 / 實際值) 的中位數
- **COD** —— 離散係數 (Coefficient of Dispersion, 越低越好)
- **PRD** —— 價格相關差異 (Price-Related Differential)

當訓練 R^2 超出 $\text{CV } R^2$ 達 0.1 以上時，會出現過擬合警示。

卡片下方有一張帶符號警示的係數表，為方便查閱而重複了 Coefficients 分頁的內容。

9.5 Coefficients (係數)

一張列出所有非零係數的表，顯示：

- 變數名
- 係數值
- 預期符號 (來自變數設定)
- 若係數方向與預期不符，則給出符號警示標記

9.6 Variable Importance (變數重要性)

標準化的係數量級：對每個預測變數計算 $|\beta| \times \text{sd}(x)$ ，因子變數則跨其虛擬變數欄彙總。有兩種顯示模式：

- 互動式 **plotly** (安裝了 **plotly** 套件時)：水平長條圖，滑鼠停留提示顯示精確的重要性、係數與相對百分比。
- 靜態 **ggplot2** (備用方案)：水平長條圖，座標軸標籤帶千位逗號。

圖表下方有一張 **DataTable**，顯示 **Variable**、**Importance**、**Coefficient** 與 **Relative %**。

9.7 Contributions (貢獻)

逐變數的偏效應圖，展示每個預測變數對模型預測值的貢獻。請從下拉式選單中選擇一個變數。

對數值型預測變數：一張該變數的 x 值對其貢獻 ($\beta \times x$) 的散布圖，附一條紅色線段和一個斜率標註。斜率標註顯示每單位的邊際效應 (例如 $+1,234.56/\text{unit}$)，對於緯度 / 經度這類取值範圍很小的變數，單位會自適應調整 (例如 $+0.12/0.001$)。副標題顯示截距 (基準) 值。

對因子變數：一張盒鬚圖，展示各因子水準上貢獻值的分布。

對雙變數交互項：三幅上下疊放的視覺化圖——聯合貢獻的三維曲面圖 (可用時為互動式 **plotly**，否則為靜態 **persp()**)、按貢獻著色的散布圖，以及跨兩個變數取值範圍的平均貢獻熱圖。

9.8 ANOVA (變異數分析)

變異數分解表，為每個預測變數顯示：

- **SS** ——平方和： $\sum (\text{contribution}_j - \overline{\text{contribution}_j})^2$
- **% of Model SS** ——佔模型總平方和的百分比
- **Coefficient** ——該預測變數的主導係數

表中含一列截距和一列 **TOTAL MODEL** (模型合計)。交互項 (名稱中含「:」者) 作為單獨的組顯示。

9.9 Diagnostics (診斷)

四幅診斷圖，採用大字級 (基準字級 16pt)、15 個座標軸刻度，標籤帶千位逗號 (不使用科學記號)：

1. 係數路徑 (Coefficient Path) ——展示隨著 $\log(\lambda)$ 增大，所有係數如何變化。曲線按變數著色。有助於直觀理解正則化路徑與變數選擇過程。
2. **CV 誤差** (CV Error) ——交叉驗證誤差隨 $\log(\lambda)$ 的變化。散點為 CV 誤差平均數；誤差槓為 ± 1 個標準誤。虛線垂直線標出 $\lambda_{\min}$ (藍色) 與 λ_{1se} (綠色)。
3. 實際值對預測值 (Actual vs Predicted) ——觀測值與預測值的散布圖。散點應聚集在 45 度虛線參照線附近。散布越寬，說明預測誤差越大。
4. 殘差對擬合值 (Residuals vs Fitted) ——殘差與擬合值的散布圖。應呈現圍繞零線 (虛線) 的隨機散布。若出現某種模式，則提示模型設定有誤 (例如非線性、異質變異數)。

每幅圖都有一個 **Download PNG** 按鈕，可按 150 DPI 儲存。

9.10 Glmnet Output (Glmnet 輸出)

原始模型輸出，顯示：所用隨機種子、模型打印輸出、選定的 λ 、 $\lambda_{\min}/\lambda_{1se}$ 、 γ (若啟用鬆弛 **lasso**)，以及完整的係數向量。

9.11 Report (報告)

設定報告欄位（估價師姓名、不動產地址、報告日期、檔案編號），並可直接從該分頁匯出為 Word (.docx) 或 PDF。完整的基於 Quarto 的報告工作流程 (HTML、Word 或 PDF) 見第 13 章。

10 下載資料

擬合之後，**Download Output (Excel)**（下載輸出，Excel）按鈕（側邊欄第 7 節，在所有用途模式下均可用）會把含預測值與診斷資訊的 Excel 檔案匯出到專案的輸出資料夾。該小節標題在估價與市場模式下為「Download Estimated Sale Prices & Residuals」（下載估計成交價格與殘差），在通用模式下為「Download Estimated Target Variable(s) & Residuals」（下載估計目標變數與殘差）。

10.1 輸出欄位

欄位	說明
est_<target>	模型預測值（例如 est_sale_price）
residual	實際值 - 預測值
cqa	狀況- 品質- 吸引力分數（Condition-Quality-Appeal，0-10 分制）
residual_sf	殘差 / 居住面積（若已指定 living_area）
cqa_sf	依據 residual_sf 排名計算出的 CQA
<var>_contribution	各變數對預測值的貢獻
basis	截距值（所有資料列相同）

成功完成後，下載按鈕上會出現一個白色勾號。

10.2 CQA 分數

CQA 把每一資料列的殘差與其餘各列相比較，按 0-10 分制排名：

- **CQA 高（~9-10）**：成交價遠高於預測值——很可能狀況 / 品質 / 吸引力更優
- **CQA 低（~0-1）**：成交價遠低於預測值——很可能狀況較差或屬不良資產處分
- **CQA ~5**：接近中位數

在估價 / 市場模式下，若指定了 living_area 欄，各資料列會按 residual_sf 遞減排序。

11 RCA 計算與下載

RCA (Residual Constraint Approach, 殘差約束法) 工作流程在 **Appraisal** (估價) 與 **Market Area Analysis** (市場區域分析) 模式下、模型擬合之後可用。後續的銷售比較表格 (第 12 章) 僅限估價模式。

11.1 開啟 RCA 對話方塊

點選側邊欄第 8 節的 **Calculate RCA Adjustments & Download** 按鈕。然後選擇：

- 分數類型：**CQA** 或 **CQA per SF** (每平方英尺 CQA)。若選擇「CQA per SF」，標的的殘差會按居住面積縮放。
- 輸入標的的 CQA 分數 (0.00–9.99，預設 5.00)

11.2 CQA 分數內插

1. 對各可比案例的 CQA 分數與殘差進行排序
2. 用線性內插把您給定的 CQA 值對應到一個殘差
3. 標的價值 = 模型預測值 + 內插所得殘差

11.3 輸出欄位

欄位	說明
subject_value	模型預測值 + 內插所得殘差
<var>_adjustment	標的貢獻 – 可比案例貢獻
residual_adjustment	標的殘差 – 可比案例殘差
net_adjustments	所有調整值之和
gross_adjustments	各調整值絕對值之和
adjusted_sale_price	可比案例成交價格 + 淨調整額

計算成功後，按鈕上會出現一個白色勾號。

12 銷售比較表格

銷售比較表格 (Sales Comparison Grid) 僅在 **Appraisal** (估價) 模式下、計算完 RCA 調整值 (第 8 步) 之後可用。它會產生一份格式化的 Excel 活頁簿，可直接編入估價報告。

12.1 可比案例選擇對話方塊

點選側邊欄第 9 節的 **Generate Sales Grid & Download** 按鈕，會彈出一個強制回應對話方塊，列出 RCA 輸出中的所有可比成交案例。可比案例分為兩組：

- **Recommended comps** (推薦的可比案例，預先勾選)：毛調整額 < 成交價格的 25%，按成交時距排序 (最新在前)。
- **Other comps** (其他可比案例，預設不勾選)：其餘的可比案例，按成交時距排序。

最多可選 30 個可比案例，然後點選 **Generate Sales Grid** 產生活頁簿。

12.2 活頁簿版面

產生的活頁簿最多含 10 個工作表，每個工作表並排放置 3 個可比案例：

- 表頭列：標的與可比案例的地址、成交價格、成交日期以及在市天數
- 分組列：位置 (緯度、經度、區域)、基地面積 (基地面積、基地尺寸) 與屋齡 (實際屋齡、有效屋齡)，帶組標題與子項目
- 模型變數列：每個非零模型係數一列，顯示標的貢獻、各可比案例貢獻以及調整百分比
- 剩餘特徵列：留空且未鎖定的儲存格，供估價師填寫模型未涵蓋的特徵 (例如狀況、品質、景觀)
- 彙總列：淨調整百分比、毛調整百分比與調整後成交價格，均帶有即時的 Excel 公式

12.3 工作表保護與剩餘特徵儲存格

每個工作表都設有密碼保護以防誤改，但剩餘特徵的取值儲存格是明確解鎖的。這樣估價師既能填入模型之外的特徵取值，又能保留由公式驅動的調整計算。

13 下載報告

13.1 兩階段 Quarto 工作流程

從側邊欄匯出報告採用兩階段的 Quarto 工作流程：

1. **Generate Quarto Report** (產生 Quarto 報告，側邊欄第 10 節) ——把一個自足的 Quarto 報告包 (.qmd 原始檔加上圖形、report_data.rds 與 reference.docx) 寫入專案的 glmnet 輸出資料夾。
2. **Convert Quarto Report** (轉換 Quarto 報告，側邊欄第 11 節) ——把 .qmd 檔案轉譯為您選擇的輸出格式。路徑預設指向最近產生的報告包，但您也可以瀏覽選擇任何一個 .qmd。

13.2 報告格式

轉換步驟中有三種格式可選：

- **HTML** ——自足的 HTML 文件。在所有平台上都能使用，除 pandoc 外無需額外軟件。推薦的預設選項。
- **Word (.docx)** ——格式化文件，便於編輯與分送。在所有平台上均可使用。
- **PDF** ——透過 Quarto 配合 LaTeX 產生。紙張尺寸遵循地區設定 (Letter 或 A4)。需要安裝 LaTeX (參見系統需求附錄)。若未偵測到 LaTeX，格式核取方塊中會自動隱藏 PDF 選項。

有 Quarto 可用時使用 Quarto 產生報告，否則改用 rmarkdown。Report 分頁另外還提供直接匯出為 Word 和匯出為 PDF 的按鈕。

13.3 報告內容

報告涵蓋所有分頁的內容：

- 資料集說明 (檔名、觀測數、反應變數、預測變數、用途)
- 模型設定 (alpha、lambda、family、standardize、是否鬆弛、非零係數)
- 結果摘要 (R^2 、調整 R^2 、廣義 R^2 、CV R^2 、RMSE、MAE)
- 模型方程式 (以 LaTeX 轉譯)
- 係數表
- 變數重要性 (長條圖與表格)
- 與 Contributions 分頁一致的逐變數貢獻圖：
 - 數值型預測變數：帶線性擬合線的散布圖
 - 因子變數：按水準分組的盒鬚圖
 - 交互項：按貢獻著色的散布圖、平均貢獻熱圖，以及靜態三維曲面快照 (persp)
- Enabled Interactions (已啟用交互項) 矩陣 (啟用了任何交互項時) ——兩個座標軸都是預測變數，納入 $x_1:x_2$ 乘積項之處打勾，以橫向頁面呈現
- 相關係數矩陣熱圖
- ANOVA 分解表
- 診斷圖：
 - 係數路徑
 - 交叉驗證誤差
 - 實際值對預測值
 - 殘差對擬合值
 - 常態 Q-Q 圖
- 模型輸出

所有座標軸標籤與色彩圖例中的數字均帶千位逗號（不使用科學記號）。
報告產生成功後，報告按鈕上會出現一個白色勾號；轉譯過程中會顯示已用時間計時器。

14 與 earthUI 的比較

earth 應用程式及其 glmnet 後處理模組同屬一套迴歸建模工具。它們共用相同的資料格式、特殊欄位型別、RCA 工作流程和示範資料集，但使用不同的建模引擎。

14.1 主要差異

特性	glmnetUI	earthUI
方法	彈性網絡 (glmnet)	MARS/Earth (earth)
變數選擇	經由 lasso 罰則自動進行	經由前向 / 後向修剪自動進行
係數關係	僅限線性效應	分段線性 (鉸鏈函數)
交互項	經由交互項矩陣兩兩組合	最高 3 階
係數符號	可透過約束強制施加	無法直接強制施加
鬆弛擬合	可使用鬆弛 lasso	不適用
Alpha 調校	可用網格搜尋	不適用 (無混合參數)
銷售比較表格	可用 (第 9 步)	可用 (第 8 步)
g 函數圖	不適用	歸組項視覺化
報告格式	HTML、Word、PDF	HTML、Word、PDF

14.2 共有的功能

兩個工具都提供：

- CSV 與 Excel 匯入，並把欄名轉換為 snake_case
- 支援 31 個國家的地區設定，含 CSV 分隔符號、小數點記號與紙張尺寸
- 特殊欄位指定 (contract_date、living_area、area、latitude、longitude 等)
- 依據基準日自動計算 sale_age
- 深色 / 淺色模式切換
- 透過 localStorage 持續保存設定
- 相關性熱圖、變數重要性、貢獻圖、ANOVA、診斷圖
- CQA 評分與 RCA 調整工作流程
- 帶可比案例選擇和格式化 Excel 輸出的銷售比較表格
- 相容同一批示範資料集 (Appraisal_1.csv)

14.3 何時該用哪一個

- 使用 glmnetUI，當您想要一個簡潔、可解釋、並能自動選擇變數的線性模型時。當預測變數很多而您希望由模型挑出最重要的那些時，彈性網絡尤其合適。
- 使用 earthUI，當您預期存在非線性關係（例如居住面積對價格的影響在不同面積區間有所不同），或需要用鉸鏈函數刻畫閾值效應時。earth 模型還能刻畫更高階的交互作用（2 階和 3 階）。此外，earthUI 還提供一項基於 Prolog/DCG 的實驗性備註擷取功能，這在 glmnetUI 中沒有對應項。
- 兩者並用，以在同一份資料上比較線性模型與非線性模型。若兩者給出的 R^2 相近，出於可解釋性考量，可能更宜採用較簡潔的 glmnet 模型。

15 示範資料集：Appraisal_1.csv

15.1 說明

示範資料集隨 glmnetUI 一同發布（並與 earthUI 共用）。可用以下方式定位：

```
demo_file <- system.file("extdata", "Appraisal_1.csv", package = "glmnetUI")
```

該檔案包含 1,502 筆住宅成交紀錄（外加第 1 列的 1 筆估價標的），取自一份模擬的 MLS 匯出檔案。資料代表一個多區域市場中的獨棟住宅成交，涵蓋不同的不動產面積、屋齡與位置。

這並非真實資料，但以北加州一個真實街廓為原型。所有身分識別資訊均已更改或移除。

15.2 各欄位

欄位	型別	特殊型別	說明
weight	numeric	weight	觀測權重（0 = 不參與擬合）
id	numeric	display_only	內部紀錄 ID
property_id	numeric	display_only	MLS 不動產識別碼
listing_id	character	display_only	MLS 掛牌編號
parcel_number	character	display_only	縣估稅官地號（APN）
street_address	character	display_only	不動產地址
city_name	character	display_only	城市
postal_code	character	display_only	郵遞區號
county_name	character	display_only	縣
contract_date	Date	contract_date	成交合約日期（用於計算 sale_age）
sale_age	numeric	—	從合約日期到基準日的天數
sale_price	numeric	（目標變數）	成交價格——反應變數
living_sqft	numeric	living_area	總居住面積（平方英尺）
beds_total	integer	—	臥室數量
baths_total	numeric	—	衛浴總數（例如 2.5 = 2 套全衛浴 + 1 套半衛浴）
lot_size	numeric	lot_size	基地面積（平方英尺）
area_id	integer	area	MLS 區域識別碼
age	numeric	actual_age	不動產屋齡（年）
latitude	numeric	latitude	緯度（入模前四捨五入到小數點後 3 位）
longitude	numeric	longitude	經度（入模前四捨五入到小數點後 3 位）
garage_spaces	integer	—	車庫車位數
days_on_market	integer	dom	在市天數
listing_date	Date	listing_date	掛牌日期
sale_concessions	numeric	concessions	賣方讓利

15.3 建議的快速上手流程

1. 啟動 glmnet 常式：earthUI::launch_glmnet()
2. 新建一個用途為 **For Appraisal** 的專案（第 1 節）
3. 把 Appraisal_1.csv 放入專案的輸入資料夾，並在第 2 節中選中它（必要時點選 **Refresh**）
4. 選擇 sale_price 作為目標變數
5. 按上表所示指定各欄位的特殊型別

6. 納入以下預測變數：sale_age、living_sqft、baths_total、lot_size、area_id（作為因子）、age、latitude、longitude、garage_spaces
7. 保持 $\alpha = 1$ (Lasso)，使用 CV 配 lambda.1se
8. 點選 **Fit Glmnet Model**
9. 檢視 Summary、Coefficients 與 Diagnostics 分頁
10. 下載輸出（第 7 步），檢視 CQA 排名
11. 以約 ~5.00 的 CQA 分數計算 RCA 調整值（第 8 步）
12. 用推薦的可比案例產生銷售比較表格（第 9 步）

16 系統需求與疑難排解

16.1 支援的平台

glmnetUI 可在 **macOS**、**Windows** 與 **Linux** 上執行。強烈建議使用 RStudio Desktop (2023.06 及以上版本)——它內建產生報告所需的 **pandoc**。

16.2 選用的相依套件

- **LaTeX** (TinyTeX、MiKTeX 或 MacTeX)：僅用於產生 PDF 報告。若未偵測到，PDF 選項會自動隱藏。安裝方式：`tinytex::install_tinytex()`
- **sysfonts + showtext**：為圖形提供 Roboto Condensed 字型。若不可用或處於離線狀態，會自動改用系統無襯線字型。
- **earth** ($\geq 5.3.0$)：僅在使用 earthUI 匯入管道時需要。

16.3 各平台說明

macOS：問題最少。安裝 TinyTeX 以產生 PDF：`tinytex::install_tinytex()`

Windows：與 RStudio 配合良好。公司配發或權限受限的機器可能對暫存目錄有限制——應用程式會在主控台給出警告，但仍會繼續執行，只是設定不能持續保存。

Linux：可能需要安裝系統程式庫：`sudo apt install libcurl4-openssl-dev libssl-dev libxml2-dev libsqlite3-dev libfontconfig1-dev`

16.4 功能降級

缺少的元件	表現
沒有 LaTeX	隱藏 PDF 選項。HTML 與 Word 仍可使用。
沒有網路 / 字型載入失敗	使用系統無襯線字型。主控台記錄相應訊息。
沒有 RSQLite / 檔案系統唯讀	設定不能持續保存。應用程式照常執行。
暫存目錄不可寫入	報告產生失敗並給出明確錯誤。請把 TMPDIR 設為可寫入的位置。

16.5 疑難排解

「PDF option not available」(PDF 選項不可用)：在 R 中執行 `tinytex::install_tinytex()`，然後重新啟動應用程式。

「Settings will not persist」(設定無法保存)：檢查使用者資料目錄的權限 (macOS：`~/Library/Application Support/R/glmnetUI/`，Linux：`~/.local/share/R/glmnetUI/`，Windows：`%APPDATA%/R/data/glmnetUI/`)。

連接埠 7879 已被佔用：執行 `lsof -ti:7879 | xargs kill` (macOS/Linux) 或 `Stop-Process -Id (Get-NetTCPConnection -LocalPort 7879).OwningProcess` (Windows PowerShell)。

參考文獻

- Friedman, Jerome, Trevor Hastie, and Robert Tibshirani (2010). "Regularization Paths for Generalized Linear Models via Coordinate Descent". In: *Journal of Statistical Software* 33.1, pp. 1–22. DOI: [10.18637/jss.v033.i01](https://doi.org/10.18637/jss.v033.i01).
- Simon, Noah et al. (2011). "Regularization Paths for Cox's Proportional Hazards Model via Coordinate Descent". In: *Journal of Statistical Software* 39.5, pp. 1–13. DOI: [10.18637/jss.v039.i05](https://doi.org/10.18637/jss.v039.i05).

Tay, J. Kenneth, Balasubramanian Narasimhan, and Trevor Hastie (2023). “Elastic Net Regularization Paths for All Generalized Linear Models”. In: *Journal of Statistical Software* 106.1, pp. 1–31. DOI: [10.18637/jss.v106.i01](https://doi.org/10.18637/jss.v106.i01).